OS 90
1204

#2

OIPE

RAW SEQUENCE LISTING   DATE: 12/17/2001
PATENT APPLICATION: US/09/842,256   TIME: 14:43:16

Input Set : **N:\Crf3\RULE60\09842256.txt**
Output Set: **N:\CRF3\12172001\I842256.raw**

SEQUENCE LISTING

```
 4 (1) GENERAL INFORMATION:
 6       (i) APPLICANT: Chambon, Pierre
 7                      Gronemeyer, Hinrich
 8                      Voegel, Johannes
 9                      Lutz, Yves
11      (ii) TITLE OF INVENTION: Transcriptional Intermediary Factor-2
13     (iii) NUMBER OF SEQUENCES: 14
15      (iv) CORRESPONDENCE ADDRESS:
16           (A) ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
17           (B) STREET: 1100 New York Avenue, NW, Suite 600
18           (C) CITY: Washington
19           (D) STATE: DC
20           (E) COUNTRY: USA
21           (F) ZIP: 20005-3934
23       (v) COMPUTER READABLE FORM:
24           (A) MEDIUM TYPE: Floppy disk
25           (B) COMPUTER: IBM PC compatible
26           (C) OPERATING SYSTEM: PC-DOS/MS-DOS
27           (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
29      (vi) CURRENT APPLICATION DATA:
30           (A) APPLICATION NUMBER: US/09/842,256
31           (B) FILING DATE: 26-Apr-2001
32           (C) CLASSIFICATION:
34     (vii) PRIOR APPLICATION DATA:
35           (A) APPLICATION NUMBER: 08/891,640
36           (B) FILING DATE:
38    (viii) ATTORNEY/AGENT INFORMATION:
39           (A) NAME: Steffe, Eric K.
40           (B) REGISTRATION NUMBER: 36,688
41           (C) REFERENCE/DOCKET NUMBER: 1383.0130001/EKS
43      (ix) TELECOMMUNICATION INFORMATION:
44           (A) TELEPHONE: 202-371-2600
45           (B) TELEFAX: 202-371-2540
48 (2) INFORMATION FOR SEQ ID NO: 1:
50       (i) SEQUENCE CHARACTERISTICS:
51           (A) LENGTH: 6156 base pairs
52           (B) TYPE: nucleic acid
53           (C) STRANDEDNESS: double
54           (D) TOPOLOGY: linear
56      (ii) MOLECULE TYPE: cDNA
59      (ix) FEATURE:
60           (A) NAME/KEY: CDS
61           (B) LOCATION: 163..4554
64      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
66 GGCGGCCGCA GCCTCGGCTA CAGCTTCGGC GGCGAAGGTC AGCGCCGACG GCAGCCGGCA        60
68 CCTGACGGCG TGACCGACCC GAGCCGATTT CTCTTGGATT TGGCTACACA CTTATAGATC       120
```

C--> 30 (at line 30)
C--> 31 (at line 31)

ENTERED

RAW SEQUENCE LISTING   DATE: 12/17/2001
PATENT APPLICATION: US/09/842,256 TIME: 14:43:16

Input Set : N:\Crf3\RULE60\09842256.txt
Output Set: N:\CRF3\12172001\I842256.raw

```
70  TTCTGCACTG TTTACAGGCA CAGTTGCTGA TATGTGTTCA AG ATG AGT GGG ATG      174
71                                                   Met Ser Gly Met
72                                                    1
74  GGA GAA AAT ACC TCT GAC CCC TCC AGG GCA GAG ACA AGA AAG CGC AAG     222
75  Gly Glu Asn Thr Ser Asp Pro Ser Arg Ala Glu Thr Arg Lys Arg Lys
76   5                  10              15                  20
78  GAA TGT CCT GAC CAA CTT GGA CCC AGC CCC AAA AGG AAC ACT GAA AAA     270
79  Glu Cys Pro Asp Gln Leu Gly Pro Ser Pro Lys Arg Asn Thr Glu Lys
80              25              30                  35
82  CGT AAT CGT GAA CAG GAA AAT AAA TAT ATA GAA GAA CTT GCA GAG TTG     318
83  Arg Asn Arg Glu Gln Glu Asn Lys Tyr Ile Glu Glu Leu Ala Glu Leu
84          40              45                  50
86  ATT TTT GCA AAT TTT AAT GAT ATA GAC AAC TTT AAC TTC AAA CCT GAC     366
87  Ile Phe Ala Asn Phe Asn Asp Ile Asp Asn Phe Asn Phe Lys Pro Asp
88          55              60                  65
90  AAA TGT GCA ATC TTA AAA GAA ACT GTG AAG CAA ATT CGT CAG ATC AAA     414
91  Lys Cys Ala Ile Leu Lys Glu Thr Val Lys Gln Ile Arg Gln Ile Lys
92      70              75                  80
94  GAA CAA GAG AAA GCA GCA GCT GCC AAC ATA GAT GAA GTG CAG AAG TCA     462
95  Glu Gln Glu Lys Ala Ala Ala Ala Asn Ile Asp Glu Val Gln Lys Ser
96  85              90                  95                  100
98  GAT GTA TCC TCT ACA GGG CAG GGT GTC ATC GAC AAG GAT GCG CTG GGG     510
99  Asp Val Ser Ser Thr Gly Gln Gly Val Ile Asp Lys Asp Ala Leu Gly
100             105             110                 115
102 CCT ATG ATG CTT GAG GCC CTT GAT GGG TTC TTC TTT GTA GTG AAC CTG     558
103 Pro Met Met Leu Glu Ala Leu Asp Gly Phe Phe Phe Val Val Asn Leu
104             120             125             130
106 GAA GGC AAC GTT GTG TTT GTG TCA GAG AAT GTG ACA CAG TAT CTA AGG     606
107 Glu Gly Asn Val Val Phe Val Ser Glu Asn Val Thr Gln Tyr Leu Arg
108             135             140             145
110 TAT AAC CAA GAA GAG CTG ATG AAC AAA AGT GTA TAT AGC ATC TTG CAT     654
111 Tyr Asn Gln Glu Glu Leu Met Asn Lys Ser Val Tyr Ser Ile Leu His
112         150             155             160
114 GTT GGG GAC CAC ACG GAA TTT GTC AAA AAC CTG CTG CCA AAG TCT ATA     702
115 Val Gly Asp His Thr Glu Phe Val Lys Asn Leu Leu Pro Lys Ser Ile
116 165             170             175                 180
118 GTA AAT GGG GGA TCT TGG TCT GGC GAA CCT CCG AGG CGG AAC AGC CAT     750
119 Val Asn Gly Gly Ser Trp Ser Gly Glu Pro Pro Arg Arg Asn Ser His
120             185             190             195
122 ACC TTC AAT TGT CGG ATG CTG GTA AAA CCT TTA CCT GAT TCA GAA GAG     798
123 Thr Phe Asn Cys Arg Met Leu Val Lys Pro Leu Pro Asp Ser Glu Glu
124             200             205             210
126 GAG GGT CAT GAT AAC CAG GAA GCT CAT CAG AAA TAT GAA ACT ATG CAG     846
127 Glu Gly His Asp Asn Gln Glu Ala His Gln Lys Tyr Glu Thr Met Gln
128             215             220             225
130 TGC TTC GCT GTC TCT CAA CCA AAG TCC ATC AAA GAA GAA GGA GAA GAT     894
131 Cys Phe Ala Val Ser Gln Pro Lys Ser Ile Lys Glu Glu Gly Glu Asp
132             230             235             240
134 TTG CAG TCC TGC TTG ATT TGC GTG GCA AGA AGA GTT CCC ATG AAG GAA     942
```

RAW SEQUENCE LISTING   DATE: 12/17/2001
PATENT APPLICATION: US/09/842,256   TIME: 14:43:16

                    Input Set : N:\Crf3\RULE60\09842256.txt
                    Output Set: N:\CRF3\12172001\I842256.raw

```
135 Leu Gln Ser Cys Leu Ile Cys Val Ala Arg Arg Val Pro Met Lys Glu
136 245                 250                 255                 260
138 AGA CCA GTT CTT CCC TCA TCA GAA AGT TTT ACT ACT CGC CAG GAT CTC      990
139 Arg Pro Val Leu Pro Ser Ser Glu Ser Phe Thr Thr Arg Gln Asp Leu
140                     265                 270                 275
142 CAA GGC AAG ATC ACG TCT CTG GAT ACC AGC ACC ATG AGA GCA GCC ATG     1038
143 Gln Gly Lys Ile Thr Ser Leu Asp Thr Ser Thr Met Arg Ala Ala Met
144                 280                 285                 290
146 AAA CCA GGC TGG GAG GAC CTG GTA AGA AGG TGT ATT CAG AAG TTC CAT     1086
147 Lys Pro Gly Trp Glu Asp Leu Val Arg Arg Cys Ile Gln Lys Phe His
148             295                 300                 305
150 GCG CAG CAT GAA GGA GAA TCT GTG TCC TAT GCT AAG AGG CAT CAT CAT     1134
151 Ala Gln His Glu Gly Glu Ser Val Ser Tyr Ala Lys Arg His His His
152         310                 315                 320
154 GAA GTA CTG AGA CAA GGA TTG GCA TTC AGT CAA ATC TAT CGT TTT TCC     1182
155 Glu Val Leu Arg Gln Gly Leu Ala Phe Ser Gln Ile Tyr Arg Phe Ser
156 325                 330                 335                 340
158 TTG TCT GAT GGC ACT CTT GTT GCT GCA CAA ACG AAG AGC AAA CTC ATC     1230
159 Leu Ser Asp Gly Thr Leu Val Ala Ala Gln Thr Lys Ser Lys Leu Ile
160                 345                 350                 355
162 CGT TCT CAG ACT ACT AAT GAA CCT CAA CTT GTA ATA TCT TTA CAT ATG     1278
163 Arg Ser Gln Thr Thr Asn Glu Pro Gln Leu Val Ile Ser Leu His Met
164             360                 365                 370
166 CTT CAC AGA GAG CAG AAT GTG TGT GTG ATG AAT CCG GAT CTG ACT GGA     1326
167 Leu His Arg Glu Gln Asn Val Cys Val Met Asn Pro Asp Leu Thr Gly
168         375                 380                 385
170 CAA ACG ATG GGG AAG CCA CTG AAT CCA ATT AGC TCT AAC AGC CCT GCC     1374
171 Gln Thr Met Gly Lys Pro Leu Asn Pro Ile Ser Ser Asn Ser Pro Ala
172         390                 395                 400
174 CAT CAG GCC CTG TGC AGT GGG AAC CCA GGT CAG GAC ATG ACC CTC AGT     1422
175 His Gln Ala Leu Cys Ser Gly Asn Pro Gly Gln Asp Met Thr Leu Ser
176 405                 410                 415                 420
178 AGC AAT ATA AAT TTT CCC ATA AAT GGC CCA AAG GAA CAA ATG GGC ATG     1470
179 Ser Asn Ile Asn Phe Pro Ile Asn Gly Pro Lys Glu Gln Met Gly Met
180                 425                 430                 435
182 CCC ATG GGC AGG TTT GGT GGT TCT GGG GGA ATG AAC CAT GTG TCA GGC     1518
183 Pro Met Gly Arg Phe Gly Gly Ser Gly Gly Met Asn His Val Ser Gly
184             440                 445                 450
186 ATG CAA GCA ACC ACT CCT CAG GGT AGT AAC TAT GCA CTC AAA ATG AAC     1566
187 Met Gln Ala Thr Thr Pro Gln Gly Ser Asn Tyr Ala Leu Lys Met Asn
188         455                 460                 465
190 AGC CCC TCA CAA AGC AGC CCT GGC ATG AAT CCA GGA CAG CCC ACC TCC     1614
191 Ser Pro Ser Gln Ser Ser Pro Gly Met Asn Pro Gly Gln Pro Thr Ser
192         470                 475                 480
194 ATG CTT TCA CCA AGG CAT CGC ATG AGC CCT GGA GTG GCT GGC AGC CCT     1662
195 Met Leu Ser Pro Arg His Arg Met Ser Pro Gly Val Ala Gly Ser Pro
196 485                 490                 495                 500
198 CGA ATC CCA CCC AGT CAG TTT TCC CCT GCA GGA AGC TTG CAT TCC CCT     1710
199 Arg Ile Pro Pro Ser Gln Phe Ser Pro Ala Gly Ser Leu His Ser Pro
```

RAW SEQUENCE LISTING   DATE: 12/17/2001
PATENT APPLICATION: US/09/842,256   TIME: 14:43:16

Input Set : N:\Crf3\RULE60\09842256.txt
Output Set: N:\CRF3\12172001\I842256.raw

```
200                505                510                515
202 GTG GGA GTT TGC AGC AGC ACA GGA AAT AGC CAT AGT TAT ACC AAC AGC    1758
203 Val Gly Val Cys Ser Ser Thr Gly Asn Ser His Ser Tyr Thr Asn Ser
204                520                525                530
206 TCC CTC AAT GCA CTT CAG GCC CTC AGC GAG GGG CAC GGG GTC TCA TTA    1806
207 Ser Leu Asn Ala Leu Gln Ala Leu Ser Glu Gly His Gly Val Ser Leu
208                535                540                545
210 GGG TCA TCG TTG GCT TCA CCA GAC CTA AAA ATG GGC AAT TTG CAA AAC    1854
211 Gly Ser Ser Leu Ala Ser Pro Asp Leu Lys Met Gly Asn Leu Gln Asn
212     550                555                560
214 TCC CCA GTT AAT ATG AAT CCT CCC CCA CTC AGC AAG ATG GGA AGC TTG    1902
215 Ser Pro Val Asn Met Asn Pro Pro Pro Leu Ser Lys Met Gly Ser Leu
216 565                570                575                580
218 GAC TCA AAA GAC TGT TTT GGA CTA TAT GGG GAG CCC TCT GAA GGT ACA    1950
219 Asp Ser Lys Asp Cys Phe Gly Leu Tyr Gly Glu Pro Ser Glu Gly Thr
220                585                590                595
222 ACT GGA CAA GCA GAG AGC AGC TGC CAT CCT GGA GAG CAA AAG GAA ACA    1998
223 Thr Gly Gln Ala Glu Ser Ser Cys His Pro Gly Glu Gln Lys Glu Thr
224                600                605                610
226 AAT GAC CCC AAC CTG CCC CCG GCC GTG AGC AGT GAG AGA GCT GAC GGG    2046
227 Asn Asp Pro Asn Leu Pro Pro Ala Val Ser Ser Glu Arg Ala Asp Gly
228                615                620                625
230 CAG AGC AGA CTG CAT GAC AGC AAA GGG CAG ACC AAA CTC CTG CAG CTG    2094
231 Gln Ser Arg Leu His Asp Ser Lys Gly Gln Thr Lys Leu Leu Gln Leu
232     630                635                640
234 CTG ACC ACC AAA TCT GAT CAG ATG GAG CCC TCG CCC TTA GCC AGC TCT    2142
235 Leu Thr Thr Lys Ser Asp Gln Met Glu Pro Ser Pro Leu Ala Ser Ser
236 645                650                655                660
238 TTG TCG GAT ACA AAC AAA GAC TCC ACA GGT AGC TTG CCT GGT TCT GGG    2190
239 Leu Ser Asp Thr Asn Lys Asp Ser Thr Gly Ser Leu Pro Gly Ser Gly
240                665                · 670                675
242 TCT ACA CAT GGA ACC TCG CTC AAG GAG AAG CAT AAA ATT TTG CAC AGA    2238
243 Ser Thr His Gly Thr Ser Leu Lys Glu Lys His Lys Ile Leu His Arg
244                680                685                690
246 CTC TTG CAG GAC AGC AGT TCC CCT GTG GAC TTG GCC AAG TTA ACA GCA    2286
247 Leu Leu Gln Asp Ser Ser Ser Pro Val Asp Leu Ala Lys Leu Thr Ala
248     695                700                705
250 GAA GCC ACA GGC AAA GAC CTG AGC CAG GAG TCC AGC AGC ACA GCT CCT    2334
251 Glu Ala Thr Gly Lys Asp Leu Ser Gln Glu Ser Ser Ser Thr Ala Pro
252     710                715                720
254 GGA TCA GAA GTG ACT ATT AAA CAA GAG CCG GTG AGC CCC AAG AAG AAA    2382
255 Gly Ser Glu Val Thr Ile Lys Gln Glu Pro Val Ser Pro Lys Lys Lys
256 725                730                735                740
258 GAG AAT GCA CTA CTT CGC TAT TTG CTA GAT AAA GAT GAT ACT AAA GAT    2430
259 Glu Asn Ala Leu Leu Arg Tyr Leu Leu Asp Lys Asp Asp Thr Lys Asp
260                745                750                755
262 ATT GGT TTA CCA GAA ATA ACC CCC AAA CTT GAG AGA CTG GAC AGT AAG    2478
263 Ile Gly Leu Pro Glu Ile Thr Pro Lys Leu Glu Arg Leu Asp Ser Lys
264                760                765                770
```

RAW SEQUENCE LISTING   DATE: 12/17/2001
PATENT APPLICATION: US/09/842,256   TIME: 14:43:16

        Input Set : N:\Crf3\RULE60\09842256.txt
        Output Set: N:\CRF3\12172001\I842256.raw

```
266 ACA GAT CCT GCC AGT AAC ACA AAA TTA ATA GCA ATG AAA ACT GAG AAG        2526
267 Thr Asp Pro Ala Ser Asn Thr Lys Leu Ile Ala Met Lys Thr Glu Lys
268         775                 780                 785
270 GAG GAG ATG AGC TTT GAG CCT GGT GAC CAG CCT GGC AGT GAG CTG GAC        2574
271 Glu Glu Met Ser Phe Glu Pro Gly Asp Gln Pro Gly Ser Glu Leu Asp
272         790                 795                 800
274 AAC TTG GAG GAG ATT TTG GAT GAT TTG CAG AAT AGT CAA TTA CCA CAG        2622
275 Asn Leu Glu Glu Ile Leu Asp Asp Leu Gln Asn Ser Gln Leu Pro Gln
276 805                 810                 815                 820
278 CTT TTC CCA GAC ACG AGG CCA GGC GCC CCT GCT GGA TCA GTT GAC AAG        2670
279 Leu Phe Pro Asp Thr Arg Pro Gly Ala Pro Ala Gly Ser Val Asp Lys
280         825                 830                 835
282 CAA GCC ATC ATC AAT GAC CTC ATG CAA CTC ACA GCT GAA AAC AGC CCT        2718
283 Gln Ala Ile Ile Asn Asp Leu Met Gln Leu Thr Ala Glu Asn Ser Pro
284         840                 845                 850
286 GTC ACA CCT GTT GGA GCC CAG AAA ACA GCA CTG CGA ATT TCA CAG AGC        2766
287 Val Thr Pro Val Gly Ala Gln Lys Thr Ala Leu Arg Ile Ser Gln Ser
288         855                 860                 865
290 ACT TTT AAT AAC CCA CGA CCA GGG CAA CTG GGC AGG TTA TTG CCA AAC        2814
291 Thr Phe Asn Asn Pro Arg Pro Gly Gln Leu Gly Arg Leu Leu Pro Asn
292         870                 875                 880
294 CAG AAT TTA CCA CTT GAC ATC ACA TTG CAA AGC CCA ACT GGT GCT GGA        2862
295 Gln Asn Leu Pro Leu Asp Ile Thr Leu Gln Ser Pro Thr Gly Ala Gly
296 885                 890                 895                 900
298 CCT TTC CCA CCA ATC AGA AAC AGT AGT CCC TAC TCA GTG ATA CCT CAG        2910
299 Pro Phe Pro Pro Ile Arg Asn Ser Ser Pro Tyr Ser Val Ile Pro Gln
300                 905                 910                 915
302 CCA GGA ATG ATG GGT AAT CAA GGG ATG ATA GGA AAC CAA GGA AAT TTA        2958
303 Pro Gly Met Met Gly Asn Gln Gly Met Ile Gly Asn Gln Gly Asn Leu
304         920                 925                 930
306 GGG AAC AGT AGC ACA GGA ATG ATT GGT AAC AGT GCT TCT CGG CCT ACT        3006
307 Gly Asn Ser Ser Thr Gly Met Ile Gly Asn Ser Ala Ser Arg Pro Thr
308         935                 940                 945
310 ATG CCA TCT GGA GAA TGG GCA CCG CAG AGT TCG GCT GTG AGA GTC ACC        3054
311 Met Pro Ser Gly Glu Trp Ala Pro Gln Ser Ser Ala Val Arg Val Thr
312         950                 955                 960
314 TGT GCT GCT ACC ACC AGT GCC ATG AAC CGG CCA GTC CAA GGA GGT ATG        3102
315 Cys Ala Ala Thr Thr Ser Ala Met Asn Arg Pro Val Gln Gly Gly Met
316 965                 970                 975                 980
318 ATT CGG AAC CCA GCA GCC AGC ATC CCC ATG AGG CCC AGC AGC CAG CCT        3150
319 Ile Arg Asn Pro Ala Ala Ser Ile Pro Met Arg Pro Ser Ser Gln Pro
320         985                 990                 995
322 GGC CAA AGA CAG ACG CTT CAG TCT CAG GTC ATG AAT ATA GGG CCA TCT        3198
323 Gly Gln Arg Gln Thr Leu Gln Ser Gln Val Met Asn Ile Gly Pro Ser
324         1000                1005                1010
326 GAA TTA GAG ATG AAC ATG GGG GGA CCT CAG TAT AGC CAA CAA CAA GCT        3246
327 Glu Leu Glu Met Asn Met Gly Gly Pro Gln Tyr Ser Gln Gln Gln Ala
328         1015                1020                1025
330 CCT CCA AAT CAG ACT GCC CCA TGG CCT GAA AGC ATC CTG CCT ATA GAC        3294
```

```
              VERIFICATION SUMMARY              DATE: 12/17/2001
              PATENT APPLICATION:  US/09/842,256   TIME: 14:43:18

              Input Set : N:\Crf3\RULE60\09842256.txt
              Output Set: N:\CRF3\12172001\I842256.raw

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:790 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=3
L:1000 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=4
L:1021 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=5
L:1046 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=6
L:1064 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=7
L:1082 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=8
L:1100 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=9
L:1163 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=12
L:1172 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:1181 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=13
L:1190 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:1199 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=14
L:1208 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14
```